

Implementing and assessing computational modeling in introductory mechanics

Marcos D. Caballero* and Michael F. Schatz†
Center for Nonlinear Science and School of Physics,
Georgia Institute of Technology, Atlanta, GA 30332

Matthew A. Kohlmyer‡
Department of Physics, North Carolina State University, Raleigh, NC 27695
(Dated: July 27, 2011)

Students taking introductory physics are rarely exposed to computational modeling. In a one-semester large lecture introductory calculus-based mechanics course at Georgia Tech, students learned to solve physics problems using the VPython programming environment. During the term 1357 students in this course solved a suite of fourteen computational modeling homework questions delivered using an online commercial course management system. Their proficiency with computational modeling was evaluated in a proctored environment using a novel central force problem. The majority of students (60.4%) successfully completed the evaluation. Analysis of erroneous student-submitted programs indicated that a small set of student errors explained why most programs failed. We discuss the design and implementation of the computational modeling homework and evaluation, the results from the evaluation and the implications for instruction in computational modeling in introductory STEM courses.

PACS numbers: 01.40.Fk, 01.40.gb

Keywords: Research in Physics Education, Curricula and evaluation, Teaching methods and strategies

I. INTRODUCTION

Computation (the use of the computer to solve numerically, simulate or visualize a physical problem) has revolutionized scientific research and engineering practice. In science and engineering, computation is considered to be as important as theory and experiment.¹ Systems that cannot be solved in closed-form are probed using computation; experiments that are impossible to perform in a lab are studied numerically.^{2,3} Yet, in sharp contrast, most introductory courses fail to introduce students to computation's problem solving powers.

Using computation in introductory physics courses has several potential benefits. Students can engage in the modeling process to make complex problems tractable. This use of computation can be leveraged to explore the generality and utility of physical principles. In a way, students are participating in work that is more representative of what they will do as professional scientists and engineers.⁴⁻⁷ When constructing simulations, students are constrained by the programming language to certain syntactic structures. Hence, they must learn to contextualize problems in a way that produces a precise representation of the physical model.^{8,9} Arguably, one of computation's key strengths lies in its utility in visualizing and animating solutions to problems. These visualizations can improve students' conceptual understanding of physics.¹⁰

We have used computation in a large enrollment introductory calculus-based mechanics course at the Georgia Institute of Technology to develop students' modeling and numerical analysis skills. We have built upon previous attempts to introduce computation in introductory physics laboratories^{11,12} by extending its usage to

other aspects of students' coursework. In particular, we have taught students to construct models that predict the motion of physical systems using the VPython programming environment.¹³ We describe the design and implementation of homework problems to develop students' computational modeling skills in a high enrollment foundational physics course (Sec. III). We also provide the first evaluation and explication of students' skills when they attempt individually to solve a novel computational problem in a proctored environment (Secs. IV–VII). We discuss implications for instructional design, considerations regarding student epistemology and the assessment of knowledge transfer as well as the broader implications of teaching computation to introductory physics students (Sec. VIII).

II. APPROACHES TO IMPLEMENTING COMPUTATION

Since the development of inexpensive modern microcomputers with visual displays, there have been a number of attempts to introduce computation into physics courses. We review these attempts by decomposing them along two dimensions (*size of intended population* and *openness of the environment*) to indicate how our approach fits with previous work.

Some have worked closely with a small number of students to develop computational models in an *open computational environment*. Historical examples include the Maryland University Project in Physics and Educational Technology,^{4,5} STELLA^{6,7} and the Berkeley BOXER project.^{8,9} Open computational environments are analogous to “user-developed” codes in scientific research. Students who learn to use an open environment have the

advantage of viewing and altering the underlying algorithm on which the computational model depends. Moreover, students might learn to develop their own models that solve new problems. It is true, however, that students must devote time and cognitive effort to learning the syntax and procedures of the programming language that the open environment supports. Ideally, it is most desirable to have students focus on developing the physical model without spending excessive time and effort on the details of constructing a working code (e.g., message handling, drawing graphics, garbage collection). It is, therefore, important to consider students' experience (or lack thereof) with computation when choosing an open computational environment.

Others have developed *closed computational environments* for use at a variety of instructional levels. These environments have been deployed in a number of settings ranging from a few students to large lecture sections. Examples of closed environments include Physlets¹⁴ and the University of Colorado's Physics Educational Technology simulations.^{15,16} Closed computational environments are analogous to "canned" codes in scientific research. Students can set up and operate the program but do not construct it; nor do they have access to the underlying model or modeling algorithm ("black box" environment). User interaction in closed computational environments is often limited to setting or adjusting parameters. Closed computational environments are useful because they typically require no programming knowledge to operate, run similarly on a variety of platforms with little more than an Internet browser and produce highly visual simulations.

It is possible for computational models created in any open environment to be used as if they were developed for a closed one. Users can be restricted (formally or informally) from viewing or altering the underlying model. Models developed using Easy Java Simulations¹⁷ (EJS) have been used in a closed manner at a variety of scales and instructional levels.^{18,19} However, all the features of the physical and computational model in an EJS simulation are available as it is an open environment. Furthermore, EJS has made authoring high quality simulations accessible to students with some (but not much) programming experience. Some have proposed teaching upper-division science majors to develop computational models using EJS.²⁰

VPython,¹³ an open computational environment, has been used to teach introductory physics students to create computational models of physical phenomena.¹¹ Typically, students write all the program statements necessary to model the physical system (e.g., creating objects, assigning variables and numerical calculations). The additional details of model construction (e.g., drawing graphics, creating windows, mouse interactions) are handled by VPython and are invisible to the students. VPython supports full three dimensional graphics and animation without the additional burden to students of learning object-oriented programming.²¹ Given its roots

in the Python programming language, VPython can be a powerful foundation for students to start to learn the tools of their science or engineering trade. Moreover, VPython is an open-source, freely available environment that is accessible to users of all major computing platforms.

The Matter & Interactions (M&I) textbook²² introduces computational modeling as an integral part of the introductory physics course. Many of the accompanying laboratory activities are written with VPython in mind and a number of lecture demonstrations are VPython programs. In the traditional implementation of M&I, the practice of constructing computational models is limited to the laboratory. In a typical lab, students work in small groups to complete a computational activity by following a guided handout. They pause periodically to check their work with other groups or their teaching assistant (TA). Students' computational modeling skills are evaluated by solving fill-in-the-blank test questions in which they must write in the VPython program statements missing from a computational model.

Our approach to teaching computation uses an open environment (in VPython) and builds on our experience with M&I to extend the computational experience beyond the laboratory. We chose to use an open environment to teach computation in order to provide students with the opportunity to look inside the computational "black-box" and alter or construct the model. Furthermore, we aimed to teach students how to develop solutions to non-analytic problems. We chose VPython (e.g., instead of Java, C or Matlab) because it has a number of helpful features for novice programmers, can be used to construct high-quality three-dimensional simulations easily and is freely available to our students. VPython is also conveniently coupled to M&I allowing us to leverage our years of experience with teaching M&I. While our implementation builds on our M&I experience, it is not limited to it. We describe our implementation philosophy in the next section.

III. DESIGN AND IMPLEMENTATION OF COMPUTATIONAL HOMEWORK

We aimed to develop an instructional strategy that helps computation permeate the course but does not require that students have previous programming experience. Furthermore, this implementation had to be easily deployable across large lecture sections; the setting in which most introductory calculus-based courses are taught. Our philosophy was that students should learn computation by altering their own lab-developed programs to solve slightly modified problems. This design philosophy was informed by what research scientists do quite often; they write a program to solve a problem and then alter that program to solve a different problem that is of interest to them. We envisioned developing computational activities that would start with guided inquiry

and exploration in the laboratory followed by independent practice on homework. Students would work with TAs in the laboratory to develop a program that solves a problem. Students would then use that program individually to solve a different problem on their homework by making any modifications that were necessary.

The class of problems that becomes available to students who have learned computation is large and diverse; we chose to focus our efforts on teaching students to apply Newton’s second law iteratively to predict motion. Students taking a typical introductory mechanics course would learn several equations to predict the motion that emphasizes kinematics, a way of describing the motion without explicitly connecting changes in the motion to forces (dynamics). These kinematic formulas are quite limited; students can only apply them to problems in which the forces are constant. This can confuse students when they are presented with a situation where such formulas do not apply.²³ Furthermore, the special case of constant force motion is usually the capstone of motion prediction in an introductory mechanics course.²⁴ By contrast, computation allows instructors to start first and foremost with Newton’s second law and emphasize its full predictive power. Students can numerically model the motion of a system as long as they are able to develop a physical model of the interactions and express it in the computational environment. The numerical integration technique used to predict motion is a simple algorithm.

As a concrete example of our design, we show a mid-semester laboratory activity and homework problem in which students modeled the gravitational interaction between two bodies. In this example, students develop a VPython program that models the motion of a craft as it orbits the Earth (Fig. 1). Students later make a number of modifications to this program to solve a new problem on their homework. This example is useful because it illustrates not only the level of sophistication we expect of students but it also illustrates the types of alterations that students are asked to make on their homework.

In groups of three, students wrote a program in the laboratory to model the motion of low-mass craft as it orbits the Earth (Fig 1). In VPython, they created the objects (lines 4–6), assigned the constants and initial conditions (lines 8-10, 12–13 and 15–17) and setup the numerical integration loop (lines 19–29). The program statements in this loop included those that calculated the net force (lines 21–23) and updated the momentum (using Newton’s second law) and position of the craft (lines 25–26). When developing their physical model, students discussed the motion of the Earth; it experiences the same (magnitude) force as the space craft. Students reasoned that its change in velocity is negligible. Hence, students did not include the motion of the Earth in their VPython program. When writing this program in the lab, students could seek help from TAs at any time. The accuracy of the students’ completed code was checked by their TAs. After completing the lab, students had written a VPython program that modeled the motion of the

```

1 from __future__ import division
2 from visual import *
3
4 craft = sphere(pos = vector(10e7,0,0), color = color.white, radius = 1e6)
5 Earth = sphere(pos = vector(0,0,0), color = color.blue, radius = 6.3e6)
6 trail = curve(color = craft.color)
7
8 G = 6.67e-11
9 mcraft = 1500
10 mEarth = 5.97e24
11
12 vcraft = vector(0,2400,0)
13 pcraft = mcraft*vcraft
14
15 t = 0
16 deltat = 60
17 tf = 365*24*60*60
18
19 while t < tf:
20
21     r = craft.pos-Earth.pos
22     rhat = r/mag(r)
23     Fgrav = -G*mEarth*mcraft/mag(r)**2*rhat
24
25     pcraft = pcraft+Fgrav*deltat
26     craft.pos = craft.pos + pcraft/mcraft*deltat
27
28     trail.append(pos = craft.pos)
29     t = t + deltat
30
31 print 'Craft final position: ', craft.pos, 'meters.'
```

FIG. 1. [Color] - Under the guidance of their TAs, students wrote the VPython program above in the laboratory. This program modeled the motion of a craft (size exaggerated for visualization) orbiting the Earth over the course of one “virtual” year. To construct this model, students must create the objects and assign their positions and sizes (lines 4–6), identify and assign the other given values and relevant initial conditions (lines 8-10, 12–13 and 15–17), calculate the net force acting on the object of interest appropriately (lines 21–23) and update the momentum and position of this object in each time step (lines 25–26).

craft moving around the Earth for any arbitrary amount of time.

In the week following the lab, students solved a computational homework problem in which they used the computational model that they had written in lab to solve a modified problem. Students were asked to alter their initial conditions to predict the position and velocity of the craft after some integration time. To solve this problem successfully, students had to identify and make changes to their initial conditions (lines 4, 6, 9–10 & 12) and integration time (line 17) In addition, students had to add an additional print statement (after line 31) to print the final velocity of the craft.

Computational homework problems were deployed using the WebAssign course management system, which facilitated the weekly grading of students’ solutions. To create the homework problem, we numerically integrated several hundred initial conditions and stored the solutions, including final quantitative and qualitative results. Each student was assigned a random set of initial conditions corresponding to a particular set of results. Ran-

domization ensured that each student received a unique realization with high probability. Students used their assigned initial conditions and wrote additional statements to answer the questions posed in the problem. Students entered numeric answers into answer blanks and selected check-boxes to answer to qualitative questions. On these weekly assignments, only students' final results were graded; their code was uploaded for verification purposes, but not graded. Grading programs for structure and syntax at this large scale requires additional work by TAs who are already charged with a number of other teaching and grading tasks. Computational homework problems were generally completed in the week that followed the associated laboratory activity.

To facilitate student success and help them learn to debug their programs, each assignment contained a *test case* – an initial problem was posed for which the solution (i.e., the results from the numerical integration) was given. When writing or altering any computer program, making programming errors (bugs) is possible. Learning to debug programs is part of learning how to develop computational models. This test case ensured that a student's program worked properly and helped to instill confidence in students who might otherwise have been uncomfortable writing VPython programs without the help of their group members or TAs. After a student checked her program against the *test case*, she completed the *grading case*, a problem without a given solution.

In keeping with our overall design philosophy, most homework problems that students solved had similar designs as the aforementioned example. In particular, students built a computational model in the laboratory and independently used that fully functioning model to solve a modified problem on their homework. On the first four homework assignments, of which the previous example is the fourth, students made few modifications to their programs; altering their initial conditions and adding a new print statement. In the next several labs, students learned to model more complicated systems (e.g., three body gravitational problem, spring dynamics with drag) while learning new algorithms such as decomposing the net force vectors into radial and tangential components. Students also learned to represent these force components as arrows in VPython. On the homework problems associated with these labs, students still used their lab-developed programs to solve new problems by changing initial conditions and representing new quantities with arrows.

The last two homework problems which students solved were not related to the laboratory; we intended to emphasize the utility of learning to predict motion using Newton's second law. To solve these problems, students wrote all the statements missing from a partially completed code to predict the motion of two interacting objects. These were interactions which students had not seen before (e.g., the anharmonic potential and Lennard-Jones interaction). In these problems, we omitted the appropriate initial conditions and the statements

that numerically integrated the equations of motion. Students had to contextualize the word problem into a programming task and produce a precise representation of the problem in the VPython programming environment. With regard to programming tasks, students had to do no more than identify and assign variables and implement the motion prediction algorithm for these two problems. A similarly designed problem was used as an evaluative assignment and is discussed in detail in Sec. IV.

IV. EVALUATING COMPUTATIONAL MODELING SKILLS

Students experienced no greater difficulty when solving computational modeling homework problems than they did when solving their analytic homework. We found no statistical difference in students' performance on their computational homework and their performance on analytic homework using a rank-sum test (Analytic 84.6% vs Computational 85.8%).²⁵ However, this result did not indicate what fraction of students were able to solve these computational homework problems without assistance. While randomizing initial conditions between each student's realization ensured that students' solutions differed with high probability, working programs could be distributed easily from student to student by email. We note that the distribution of students' programs might not be deleterious; students who receive these programs must still read and interpret the program statements to enter in their initial conditions, make changes to the force law or print additional quantities. This is a more complex interaction than simply plugging numbers into an algebraic solution that they discovered online. In a sense, students who work with shared code are using a "closed" computational environment.

Nevertheless, we wanted to measure how effective students were at individually solving computational problems. We delivered a proctored laboratory assignment during the last lab of three different semesters to evaluate students' computational skills on an individual basis. Students received a partially completed program that created two objects (one low-mass and one high-mass), initialized some constants and defined the numerical integration loop structure. We aimed to evaluate students' engagement of the modeling process by contextualizing a physics problem into programming task. Furthermore, certain programming skills were being assessed, namely, students' abilities to identify and assign variables and implement the numerical integration algorithm. The assignment was delivered using WebAssign in a timed mode (30 minutes), and TAs were not permitted to help students debug their programs. A timed assignment opens with a pop-up dialog box that informs the student of the time limit. After the student acknowledges the limit (by clicking the "OK" button) the full assignment opens with a countdown clock in the upper corner of the browser window. When the time runs out, the answers that have been

TABLE I. As part of a final proctored lab assignment, students completed a partially constructed program that modeled the motion of an object under the influence of a central force. The partially written program defined the objects, some constants and the numerical integration loop structure. Delivered initial conditions, the sign (\pm) and distance dependence (r^n) of the force and object names were randomized on a per student basis. Slightly modified versions (Ver.) of this assignment were given at the end of three different semesters. Modifications were made to streamline delivery (Version 1 to Version 2), minimize transcription errors and improve presentation (Version 2 to Version 3). Students' performance on Version 1 was likely inflated because some students were allowed to work the problem on two separate occasions.

Ver.	Correct	Incorrect	% Correct
1	303	168	64.3
2	201	193	51.0
3	316	176	64.2
Overall	820	537	60.4

selected or entered are automatically submitted and the student is locked out of the assignment. The format of the assignment was identical to students' final two homework problems; students were given a test case to check their solution before solving the grading case.

For this assignment, students modeled the motion of the low-mass object as it interacted with the high-mass object through a central force. The nature of the force (attractive or repulsive) and its distance dependence (r^n) were randomized on a per student basis. We also randomized some of the variable names in the partially completed program to hinder copying. After adding and modifying the necessary program statements, students ran their program and reported the final location and velocity of the low-mass object. During the assignment, students did not receive feedback from the WebAssign system about the correctness of their solution, but they were given three attempts to enter their answers. Similar to students' online homework, only the final numerical answer was graded.

Performance varied from semester to semester (Table I) because the assignment was modified slightly between each semester in order to streamline delivery (Version 1 to Version 2), reduce transcription errors and improve presentation (Version 2 to Version 3). In the first semester, students were permitted to attempt Version 1 of the assignment twice due to a logistical issue with the initial administration of the assignment. The majority of students (64.3%) were able to model the grading case successfully on the second administration of the assignment. Students' performance on Version 1 was likely inflated because some students were able to work the problem twice.²⁶ Students solved Version 2 only once, and student performance dropped. A number of students were confused by the randomized exponent on the units of one of their initial conditions (Sec. VI). About half of the

TABLE II. Incorrectly written programs were subjected to an analysis using a set of codes developed from common student mistakes. The codes focused on three procedural areas: *using the correct given values* (IC), *implementing the force calculation* (FC) and *updating with the Newton's second law* (SL). We reviewed each of the incorrectly written student programs for each of the features listed below. These codes are explained in detail in Appendix A.

Using the correct given values (IC)	
IC1	Used all correct given values from grading case
IC2	Used all correct given values from test case
IC3	Used the correct integration time from either the grading case or test case
IC4	Used mixed initial conditions
IC5	Exponent confusion with k (interaction constant)
Implementing the force calculation (FC)	
FC1	Force calculation was correct
FC2	Force calculation was incorrect but the calculation procedure was evident
FC3	Attempted to raise separation vector to a power
FC4	Direction of the force was reversed
FC5	Other force direction confusion
Updating with Newton's second law (SL)	
SL1	Newton's second law (N2) was correct
SL2	Incorrect N2 but in an update form
SL3	Incorrect N2 attempted update with scalar force
SL4	Created new variable for \vec{p}_f
Other (O)	
O1	Attempted to update (force/momentum/position) for the massive particle
O2	Did not attempt the problem

students (51.0%) were able to model the grading case successfully. Students were more successful on Version 3 of the assignment; 64.2% modeled the grading case correctly.

Overall, roughly 40% of the students were unable to model the grading case. To determine exactly what challenges they faced while completing this assignment, we reviewed the program of each student who failed to model the grading case. Through a $\sim 60\%$ bonus on the proctored assignment, we encouraged all students to upload their programs to the WebAssign system. We limited our review to the programs submitted for Versions 2 and 3 of the assignment.

V. SYSTEMATICALLY UNFOLDING STUDENTS' ERRORS

Students must to perform several tasks to successfully write and execute the program for the proctored assignment. Students must interpret the problem statement; that is, they must contextualize a word problem into a programming task. They must review the partially completed program and identify the variables to update. Students need to apply their knowledge of predicting motion

using VPython to the problem. They must identify that the force is non-constant and then write the appropriate programming statements to calculate the vector force. Students need to then complete the motion prediction routine by writing a statement to update the momentum of the low-mass object.

Using an iterative-design approach, we developed a set of binary (affirmative/negative) codes to check which tasks students performed correctly and which errors they made. An initial review of students' uploaded programs yielded the mistakes that were made most often. These common mistakes formed the basis for the codes. The codes were developed empirically and several iterations were made before they were finalized. Two raters tested the codes by coding a single section of student submitted programs ($N = 45$). The raters resolved their differences which further explicated the codes and then re-coded the section. The final codes (Table II) were used by both raters independently to code the remaining sections ($N = 324$). The final codes had high inter-rater reliability; both raters agreed on 91% of the codes.

We classified the codes into one of three procedural areas: *using the correct given values* (IC), *implementing the force calculation* (FC) and *updating with the Newton's second law* (SL). These areas were congruent with the broad range of difficulties which students exhibited through their erroneous programs. Each code is explained in greater detail in Appendix A.

Determining where students encountered difficulties with these tasks might help explain how students learn this algorithmic approach to use Newton's second law to predict motion. Because we reviewed students' programs after they were written, we are unable to comment directly on students' challenges with contextualizing the problem. Our work was limited to analyzing students' procedural efforts (i.e., identifying variables and implementing the numerical integration algorithm). However, some information about students' thoughts and actions could be inferred from this analysis.

VI. FREQUENCY OF ERRORS IN STUDENTS' PROGRAMS

We measured the frequency of students' errors within each category (IC, FC and MP) by mapping binary patterns extracted from our coding scheme to common student mistakes. The number of possible binary patterns that we could observe in our data ranged from nine for MP to seventeen for FC with 13 possible for IC. Not all the codes within a given category are independent, hence, the number of possible binary patterns is much less than 2^n . Within a given category, we found that a large percentage of students could be characterized by just a few error patterns (between four and seven).

The errors we observed were not necessarily unique to computational problems. The most notable errors involved calculating forces or updating the momentum.

TABLE III. Only seven of the fourteen distinct code patterns for the IC category (Table II) were populated by more than 3% of the students. The patterns (ICx) are given by affirmatives (Y) and negatives (blank) in the code columns (IC#). The percentage of students with each pattern is indicated by the last column (%). These 7 patterns accounted for 88.8% of students with erroneous programs.

Initial Condition Codes						
Pattern	IC1	IC2	IC3	IC4	IC5	%
ICa		Y	Y			27.6
ICb			Y	Y		16.0
ICc	Y		Y			14.4
ICd			Y		Y	13.8
ICe				Y		7.9
ICf			Y			5.2
ICg	Y					3.8

Most of these errors appeared to be physics errors reminiscent of those made on pencil and paper problems. Many of them could have been mitigated by qualitative analysis. Some errors were unique to computational models and the iterative description of motion because they could produce a program that ran but did not model the system appropriately. Still others (e.g., replacing initial conditions) appeared to be simple careless mistakes, but, when investigated, highlighted the fragility of students' knowledge.

A. Initial Condition Errors

Students had to identify and update a total of eight given values: the interaction constant (k), the "interaction strength" (n), the mass of the less massive particle, the position and velocity of both particles and the integration time. Most students with incorrect programs (88.8%) fell into one of seven IC patterns (Table III). Students in ICa (27.6%) identified and correctly replaced all the initial conditions with those from the test case (IC2), including the integration time (IC3). Those in ICb (16.0%) mixed up the initial conditions (IC4), but used the correct integration time (IC3). Students in ICc (14.4%) identified and correctly replaced all the initial conditions with those from the grading case (IC1), including the integration time (IC3). Students who appeared in ICd (13.8%) were confused by the exponent on the units of the interaction constant (IC5), but used the correct integration time (IC3). Students in ICe (7.9%) used a variety of initial conditions and given values (IC4). Those students in ICf (5.2%) and ICg (3.8%) used incorrect initial conditions (IC3) or the wrong integration time (IC1), respectively. Most students might have simply forgotten to update one or more of the initial conditions from either the default case or the test case (ICb, ICe, ICf and ICg). A small fraction of students with mixed initial conditions had values from all three cases.

Students in ICa were most likely stuck on the test case because they had trouble with another aspect of the problem. These students were unable to obtain the solutions provided in the test case and kept working on it. It is possible a number of these students ran out of time while trying to debug their programs.

It is difficult to say definitively if students with mixed initial conditions (ICb and ICe) were unable to identify the appropriate values, as we reviewed students' programs only after they were submitted. It is possible that these students were just careless when making changes, but they might have been unable to identify and update these quantities. Some students could have been in the process of updating these quantities when they ran out of time and uploaded their programs.

Identifying and updating variables in a program is not a trivial task for students. In fact, their challenges with updating variables highlights the fragility of their computational knowledge. As an example, consider the students who confused the exponent on the length unit of the interaction constant (k) for the exponent in scientific notation of k when they defined it in their programs (ICd). The distance dependence of the central force was randomized, and hence the units of the interaction constant (k) were dependent on a student's realization. In Version 2 of the assignment, the exponent on the length unit of k was colored red (WebAssign's default behavior for random values). A student in ICd would read $k = 0.1 \text{ Nm}^3$ to mean $k = 100$ rather than $k = 0.1 \text{ Newton times meters cubed}$. In Version 3 of the assignment, we changed the exponent's text color to black like the rest of the non-random text. The overall frequency of this mistake dropped from 30.5% to 9.1%.

B. Force Calculation Errors

Students were given the magnitude of the force as an equation ($F = kr^n$) and told that their (attractive or repulsive) force acted along the line that connected the two objects. In solving this problem, students had to correctly calculate the magnitude of the central force and identify the unit vector (\hat{r}) and sign (\pm) for their own realization. Almost all students (98.8%) appeared with one of five FC patterns (Table IV). Students in FCa (23.9%) implemented the force calculation algorithm correctly (FC2), but reversed the direction of the net force (FC4). Those in FCb (22.2%) performed the force calculation correctly (FC1). Students in FCc (15.7%) implemented the procedure correctly (FC2) but were likely to include a force irrelevant to the problem (i.e., gravitational or electric interactions) or compute only the magnitude of the net force. Students who appeared in FCd (14.6%) attempted to raise the separation vector to a power. Students in FCe (14.0%) showed no evidence of an appropriate force calculation procedure; the procedure was either completely incorrect (e.g., used the differential form of the Impulse-momentum theorem) or was

TABLE IV. Only six of the nine distinct code patterns for the FC category (Table II) were populated by more than 3% of the students. The patterns (FCx) are given by affirmatives (Y) and negatives (blank) in the code columns (FC#). The percentage of students with each pattern is indicated by the last column (%). These 6 patterns accounted for 98.8% of students with erroneous programs.

Force Calculation Codes						
Pattern	FC1	FC2	FC3	FC4	FC5	%
FCa		Y		Y		23.9
FCb	Y					22.2
FCc		Y				15.7
FCd		Y	Y			14.6
FCe						14.0
FCf		Y			Y	8.4

calculated outside the numerical integration loop (i.e., a constant force). Those students in FCf (8.4%) had an appropriate force calculation procedure (FC2) but invented a unit vector for the net force (FC5).

The difficulties that students' faced when numerically computing the net force could stem from a weak grasp of the concept of vectors. Students in FCa made directional mistakes (e.g., changing the sign of one of lines 21–23 in Fig. 1) that could have been easily identified and rectified by drawing a sketch of the situation, a problem-solving strategy that is practiced in the laboratory. Those who raised the separation vector to a power (FCd) likely transcribed the central force equation (replacing r by \vec{r}) without thinking that this operation was mathematically impossible ($\vec{F} \sim k(\vec{r})^n$ vs. $\vec{F} \sim k|\vec{r}|^n\hat{r}$). We have found that students attempt a similar operation on pencil and paper problems; raising components of a vector to a power (e.g., $(\vec{r})^n = \langle r_x^n, r_y^n, r_z^n \rangle$). However, in the pencil and paper case, students are not immediately directed to their mistake as they are in a programming environment. VPython raised an exception error when this operation was attempted. These students appeared to be unable to parse this error into any useful information. Students who make this type of error might be helped by additional exposure to translating force equations to precise programmatic representations.⁹ Some students invented a unit vector (FCf) for the net force. This was most likely because they had computed a scalar force and tried to add a scalar impulse to the vector momentum. VPython raised a different exception error if an attempt to add a scalar to a vector was made. These students were able to parse this error, but resolved it incorrectly.

Other students (FCc) might have incorrectly contextualized the problem by including an irrelevant force (i.e., gravitational or electric interactions). The problem clearly stated that the two objects were far from all other objects. It did not explicitly state to neglect the gravitational interaction between the objects. However, the gravitational interaction could be safely neglected for the range of masses and distances we had chosen. Fur-

thermore, nothing about the charge of the objects was mentioned in the problem statement. It is surprising that students included these interactions in their models. One possible explanation for the inclusion of these interactions is that students had memorized how to solve the gravitational and Coulomb problems because these problems had appeared on their homework several times and on an exam. They might have panicked and simply wrote all possible forces they could remember.

A number of students (FCe) did not employ the force calculation algorithm at all. Some of these students computed the net force (e.g., lines 21–23 in Fig. 1) outside the numerical integration loop (e.g., before line 19 in Fig. 1). In this case, the net force was effectively constant and therefore only correct at $t = 0$. A program with correct syntax will run regardless of the physical implications. This error is unique to computational problems in which motion is predicted iteratively. Students in introductory physics rarely use Newton’s second law to predict motion due to non-constant forces. Other students who fell into FCe wrote “creative” program statements. Students in this group manipulated some quantities in the loop but did not perform any physically relevant calculations. The number of students with “creative” program statements was relatively small.

C. Newton’s Second Law Errors

Students had to write a program statement similar to line 25 in Fig. 1 to properly update the momentum using Newton’s second law. Most students demonstrated no difficulty in remembering the formula for the momentum update but some met challenges with making that description precise.⁹ Nearly all students (95.7%) fell into one of four SL patterns (Table V). Most students appeared in SLa (69.7%) because they wrote the momentum update correctly (SL1). A much smaller number of students fell into SLb (13.2%) and attempted to update the vector momentum with a scalar force. Students in SLc (7.9%) were unable to write Newton’s second law in any form that updated (all codes negative). A small fraction (SLd, 4.9%) wrote Newton’s second law in an iterative form, but did so incorrectly (SL2).

Students who attempted to update the momentum with a scalar force (SLb) might still face difficulties with understanding vectors. The momentum update is presented as a vector equation ($\vec{p}_f = \vec{p}_i + \vec{F}\Delta t$). These students might be unable to unpack that representation into a precise programmatic description, but it was more likely that they calculated a scalar force (FCc) and then simply wrote the correct (vector) second law syntax. VPython raised an exception error if an attempt to add a vector to a scalar was made. The students appeared unable to parse this error into any useful information.

Students who were unable to write Newton’s second law in any form that updated (SLc) might have experienced difficulties with converting the second law for-

TABLE V. Only four of the nine distinct code patterns for the SL category (Table II) were populated by more than 3% of the students. The patterns (SLx) are given by affirmatives (Y) and negatives (blank) in the code columns (SL#). The percentage of students with each pattern is indicated by the last column (%). These 4 patterns accounted for 95.7% of students with erroneous programs.

Second Law Codes					
Pattern	SL1	SL2	SL3	SL4	%
SLa	Y				69.7
SLb		Y	Y		13.2
SLc					7.9
SLd		Y			4.9

mula into a precise and useful programmatic representation. Students in this category either wrote Newton’s second law in a non-update form (e.g. writing `deltap = Fnet*deltat` or `pf - pi = Fnet*deltat` as line 25 in Fig. 1) or wrote a number of program statements that manipulated quantities but performed no useful calculations. In either case, these students could benefit from the precision required by a programming language.⁹ By forcing them to accurately represent Newton’s second law in their programs, they might begin to distinguish between the utility and applicability of its various algebraic forms.

Students who wrote Newton’s second law in form that updated incorrectly (SLd) either remembered the formula for the second law incorrectly or made a typo. These students would generally leave off the time step in the momentum update (e.g., $\mathbf{p} = \mathbf{p} + \mathbf{F}$) or divide by it ($\mathbf{p} = \mathbf{p} + \mathbf{F}/\text{deltat}$). Dividing by the time step is a particularly egregious error because the it was quite small. Hence, the impulse added in this case would be large. Students who made this error were unable to assess the state of the visualization (the particle flew off to “infinity”) to debug this error.

VII. COMMON ERROR PATTERNS IN STUDENTS’ PROGRAMS

The patterns within individual categories (IC, FC and SL) indicated the frequency of common mistakes students made when solving the proctored assignment, but a single student could make one or more of these mistakes. Evaluating a student’s complete solution requires an analysis using all the codes (Table II). In principle, the codes we developed could have up to ~ 4300 possible error patterns using all sixteen codes. In fact, the intersections of code categories indicated that the number of distinct errors made by students across all categories was relatively small; we found only 111 distinct binary patterns. It is possible to relate these unique patterns in a manner that suggests dominant common errors.

Cluster analysis, a technique borrowed from data min-

ing, is particularly well suited for this application because it characterizes patterns in complex data sets.^{27,28} This technique has been used previously to classify students' responses to questions about acceleration and velocity in two dimensions.²⁹ It was used here to determine the major features in students' incorrect programs which were responsible for their failure.

We applied the cluster analysis technique to the data generated from our set of binary codes. We used the Jaccard metric³⁰ to measure inter-cluster distances and linked clusters using their average separation.³¹ We tested several other metrics (e.g., Hamming, city block, etc.). The Jaccard metric was chosen because it neglects negative code pairs. Both the Hamming and city block metrics produced similar pairings at low levels, but higher order clusters were difficult to interpret. We used average linkage to avoid the effects of "chaining" that appeared when nearest³² and because useful clusters were more difficult to distinguish when farthest³³ neighbor linkage was used.

Thirty clusters with inter-cluster distances below 0.5 were reviewed in detail. This cutoff was selected to minimize the number of unique clusters while still rendering clusters with useful interpretations. Most students (86.5%) appeared in seven of the thirty clusters (Table VI). These clusters had very few students (<1%) with affirmatives in the "Other" category. Codes O1 and O2 were dropped from Table VI for this reason. Each of the other 23 clusters were populated by less than 3% ($N \approx 10$) of the students, and the bottom 18 clusters had less than 1% ($N \approx 3$) each. Each of the dominant clusters demonstrated a unique challenge that students faced while solving the proctored assignment (Table VI).

Students in cluster A (23.8%) tended to remain stuck on the test case (ICa) due to an error in their force calculation. Reversing the direction of the force (FCa) was the most common mistake, followed by raising the separation vector to a power (FCd). Most students in this cluster had no trouble expressing Newton's second law (SLa). These students worked diligently to solve the test case but were unable to do so. As a result, they did not proceed to the grading case.

Cluster B (19.8%) contained students who made mistakes while replacing the given values and initial conditions (any IC code except ICa). Some of these students worked with the grading case (ICc and ICg). Others might have been working with either case and had mixed conditions (ICb and ICe) or simply incorrect ones (ICf). Still others might have incorrectly assigned the exponent on the units of k to the value of k (ICd). At any rate, most students in this cluster were able to construct a working albeit incorrect program. Given their unfamiliarity with general central force interactions, these students might have believed their solutions were correct. In fact, it is possible that students who were working with the grading case (ICc and ICg) had solved the test case correctly and simply made a typo.

Students in cluster C worked with either the grading or

test case and might have made a number of mistakes with their initial conditions (any IC code except ICa). The dominant error in cluster C were students who computed the magnitude of the net force (FCc) and attempted to update the vector momentum with this scalar force (SLb). This mathematically impossible operation would have raised a VPython error. Students in this cluster were unable to parse this error into any useful information.

Cluster D (10.8%), like cluster A, was populated by students who tended to make errors in the force calculation (FCa and FCd), but students in Cluster D worked with the grading case (ICc). The most common error in Cluster D was reversing the direction of the net force (FCa) followed by raising the separation vector to a power (FCd). Again, like cluster A, most students met no challenges when updating the momentum using Newton's second law (SLb). These students might have started working with the test case, but we think it is more likely that they jumped right into working with the grading case because the dominant error appears in their force calculations.

Students in cluster E (7.6%) tended to raise the separation vector to a power (FCd) and have mixed initial conditions (ICb, ICd, ICe and ICf). These students generally had no difficulty with writing Newton's second law correctly (SLa). The dominant error for students in cluster E was raising the separation vector to a power (FCd). This mathematically impossible operation would have raised a VPython error. Students in this cluster were unable to parse this error into any useful information.

Cluster F (7.1%) contained students who worked solely with the test case (ICa) and either had no issue with their force calculation (FCb) or had no evident force calculation procedure (FCe). Most of these students had no difficulty updating the momentum using Newton's second law (SLa). Students in cluster F were able to construct a program which ran without raising any VPython errors. Students who had no issue with their solution likely completed test case but simply ran out of time before turning to the grading case. Students with no evident procedure generally computed the net force outside the numerical calculation loop, essentially making this force constant in time. Given students' unfamiliarity with general central force interactions, it would not be surprising if students who treated the central force outside the loop believed their solutions were correct.

Students in cluster G (4.1%) all invented an incorrect unit vector for the force rather than using \hat{r} (FCf) regardless of the case with which they worked (ICc, ICb and ICf). These students generally had no difficulty updating the momentum using Newton's second law (SLa). Most likely, these students computed the magnitude of the force, similar to students in cluster C, but were able to parse the resulting VPython error. Students in cluster G corrected their mistake by assigning some unit vector to the force before the momentum was updated.

TABLE VI. Only seven of the thirty clusters with an inter-cluster distance of less than 0.3 were populated by more than 3% of the students. The bottom 18 clusters were populated by less than 1% of students each. These seven clusters accounted for 86.5% of students. The percentage of affirmatives for each code (Table II) within any given cluster (A-G) is given to the nearest whole percentage. Codes with affirmative percentages greater than 60% are bolded. These clusters had very few students (< 1%) with any affirmatives in the ‘Other’ category, hence the results from this category are not reported. The percentage of students in each cluster is indicated in the last column (%).

Cluster	Initial Conditions					Force Calculation					Second Law				%
	IC1	IC2	IC3	IC4	IC5	FC1	FC2	FC3	FC4	FC5	SL1	SL2	SL3	SL4	
A	00	68	93	18	15	00	100	22	66	09	95	00	00	01	23.8
B	21	01	86	37	41	88	00	00	00	00	97	00	00	00	19.8
C	04	33	76	31	22	00	94	00	08	00	00	98	98	08	13.3
D	98	00	85	00	00	00	85	18	50	00	98	00	00	00	10.8
E	00	00	57	75	36	00	100	79	00	04	89	00	00	00	7.6
F	00	100	96	00	00	65	00	00	00	00	73	19	00	04	7.1
G	27	00	93	53	07	00	100	00	07	100	93	00	00	00	4.1

VIII. CLOSING REMARKS

Students can develop the skills necessary to predict the motion of sundry dynamical systems in large introductory physics courses. After a solving a suite of computational homework problems, most students ($\sim 60\%$) were able to model the motion of a novel problem successfully. In our work, we discovered that most students who were unsuccessful encountered challenges when calculating the net force acting on the object in the motion prediction algorithm (Clusters A and C through G in Table VI). By contrast, there were fewer students whose primary challenge was identifying and assigning variables (Cluster B in Table VI). We acknowledge that we have limited the development of our students’ computational skill set to contextualizing a word problem into a programming task, identifying and updating input variables and applying a motion prediction algorithm. We believe that further development of our homework problems and other novel deployments could broaden the scope of the skills students develop.

Procedural errors such as those we have documented (Secs. VI & VII) could be corrected through additional materials aimed at addressing each error in turn. However, the results from this work indicate that instructional efforts should be focused not only on correcting procedural mistakes but also on developing students’ qualitative habits of mind. Training students to write programs to predict motion might help them to be successful in a highly structured environment, but they would be better served by learning the practice of debugging. Here, debugging includes identifying syntax errors, of which we found few, and, more importantly, performing the type of qualitative analysis that is typically taught for solving analytic problems. Students who could synthesize their analytic and computational skills would be better prepared to solve the open-ended problems they will face in their future work.

Developing the materials to teach these skills requires

an evaluation of how students contextualize computational problems. We do not claim to understand this presently, although we have been able to glean some suggestive information based on students’ errors. Investigating what students think about when solving computational problems requires structured student interviews (i.e., a think-aloud study). In the future, we plan to perform such a study to not only characterize students’ abilities to contextualize but also to elucidate the mechanism for some of the errors we reported in Secs. VI & VII.

Research into skill development in math and science has shown a strong correlations with student epistemology.^{34,35} Epistemology is important because the views that students hold affect how they learn³⁶ and, ultimately, how successful they are in their science courses.^{37,38} It is therefore crucial that we understand students’ sentiments about learning a new tool such as computation. Our students expressed anxiety and demonstrated a lack of self confidence, even with their additional exposure to computational problems. We are developing an attitudinal survey aimed at exploring these and other beliefs in detail. Students who learn to use computational modeling and are confident in their abilities will be better prepared to solve challenging problems.

We have not claimed to have assessed a transfer of computational knowledge. We designed a set of problems (Sec. III) that students solved over the course of the semester with an eye towards a final assessment of their skills using a novel problem. This problem (Sec. IV) was similar to some of the homework problems students had solved previously. It focused on key skills that we desired students to acquire: contextualizing a problem, identifying and assigning variables in a program and carrying out the motion prediction algorithm. An evaluation of transfer would require that students apply these computational skills to a different domain (e.g., electromagnetism) or a different task (e.g., open-ended inquiry). Demonstrating transfer of computational knowledge is a necessary step in developing students into flexible prob-

lem solvers for the 21st century.

It is the goal of many reforms in physics education to develop students into flexible problem solvers while exploring the practice of science. Teaching computational modeling alongside physics provides support for that effort. Students learn the tools for doing science while developing a qualitative understanding of physical systems, exploring the generality of physics principles and learning broadly applicable problem solving methods in computation.

ACKNOWLEDGMENTS

We would like to thank Edwin Greco for his help with the deployment of our computational problems and Balachandra Suri for his help in analyzing some of the student written code. This work was supported by National Science Foundation's Division of Undergraduate Education (DUE0618519 and DUE0942076).

Appendix A: More details on the evaluation codes

The codes shown in Table II were developed empirically. The procedure followed an iterative-design approach. We reviewed student work for common errors and devised a rough coding scheme. We then tested the scheme on a new set of student submitted programs. The scheme was refined and re-tested. This iterative procedure was repeated several times until we captured the majority of students' mistakes. Each code is explained in detail below.

1. Using the correct given values (IC) Codes

We reviewed the variables in each student's program. The default values had to be updated with the values given in the problem statement in the partially completed program. We present the codes used to categorize each student's program with respect to identifying and updating the appropriate initial conditions for their realization.

IC1 – Student used all the correct given values from grading case. A student must replace the values of all the variables (mass, position, and velocity, interaction constant k and the exponent in the force law n in $F = kr^n$) with those given in the grading case. This code excluded the integration time. It was intended that the larger mass object was to remain at its location. This was made explicit in the problem statement; the initial position $\langle 5, 4, 0 \rangle$ m and velocity $\langle 0, 0, 0 \rangle$ m/s of the larger mass of object were given in the problem statement, even though these same values appeared in the partially written program.

IC2 – Student used all the correct given values from test case. A student must replace the values of all the

variables (mass, position, and velocity, interaction constant k and the exponent in the force law n in $F = kr^n$) with those given in the test case. This code excluded the integration time. It was intended that the larger mass object was to remain at its location (See IC1).

IC3 – Student used the correct integration time from either the grading case or test case. A student must replace the default integration time (1 s) with the values given in the case with which they intended to work (grading or test). A student who mixed initial conditions was given an affirmative on this code if the majority of their initial conditions were from the same case as the integration time.

IC4 – Student used mixed initial conditions. A student who used some but not all of the initial conditions from any of the cases (default, test, or grading) was given an affirmative on this code. This code excluded the integration time.

IC5 – Students confused the exponents on the units the exponent of k (interaction constant). Many students incorrectly thought the exponent on the length unit of the interaction constant was the scientific notation exponent for the interaction constant itself. For example, a student thought $k = 0.1 \text{ Nm}^3$ meant $k = 100$ rather than $k = 0.1 \text{ Newton times meters cubed}$.

2. Implementing the force calculation (FC) Codes

We reviewed how the students employed the force calculation algorithm in each of the programs written for the proctored assignment. The partially written program given to the students left out all statements related to the force calculation. Students were required to fill in this procedure using the appropriate VPython syntax. We present the codes used to categorize each student's program with respect to computing the vector force acting on the low-mass object.

FC1 – The force calculation was correct. A student must compute the separation vector, its magnitude, its unit vector, the magnitude of the force and the vector force correctly. Each of these steps may be combined as long as the final result computes the vector force acting on the less massive particle at each instant. These steps must all appear in the numerical integration loop.

FC2 – The force calculation was incorrect, but the calculation procedure was evident. In the numerical integration loop, the student must perform a position vector subtraction, a calculation of the force magnitude and some attempt at combining magnitude with unit vector (any unit vector was acceptable). If a student treated the problem using components and had some force which is a vector, it was coded as evident. If any part of the calculation was performed outside the loop, it was coded as not evident.

FC3 – The student attempted to raise the separation vector (\vec{r}) to a power. Students who raised the separation vector to a power generated a VPython exception

error:

```
unsupported operand type(s) for ** or pow():
'vector' and 'int'.
```

This error told them that VPython cannot raise a vector to a power, as it is a mathematically impossible operation.

FC4 – The direction of the force was reversed. Students had to assign the correct unit vector and sign to the force depending on whether their force was attractive or repulsive. This code was not used if the student calculated the force as a magnitude only, raised \vec{r} to a power, or invented a unit vector (e.g., $\langle 1, 0, 0 \rangle$). Visual feedback (i.e., the lower mass particle flying off to infinity) indicated a simple sign mistake.

FC5 – Student had some other force direction confusion. Some students used vectors other than \vec{r} or $-\vec{r}$ to compute \vec{F} . Other students computed the force as a magnitude and then multiplied it by an “invented” unit vector (e.g., $\langle 1, 0, 0 \rangle$, \hat{p}). Both of these errors were given an affirmative for this code.

3. Updating with the Newton’s second law (SL) Codes

We reviewed how the students employed the momentum update in each of the programs written for the proc-tored assignment. The partially written program given to the students left out the one line of code necessary to update the momentum. Students were required to fill in this line using the appropriate VPython syntax. We present the codes used to categorize each student’s program with respect to updating the momentum of the low-mass particle.

SL1 – Newton’s second law was correct. Correct Newton’s second law meant that it was “correct as a physics principle” and also that it appeared “in the update form”. This meant that `pfinal = pinitial + Fnet*deltat` alone in a loop did not fall under “correct Newton’s second law”. It is an incorrect update form.

SL2 – Newton’s second law was incorrect but in form that updates. Newton’s second law updates the momentum, but not necessarily correctly. (e.g., `p = p + Fnet`, `p = p + Fnet/dt`, `pf = p + Fnet`, etc.)

SL3 – Newton’s second law was incorrect and the student attempted to update it with a scalar force. Some students computed the magnitude of the force acting on the particle and then used this magnitude to update the momentum. Students who did this raised a VPython exception error:

```
unsupported operand type(s) for +: 'vector'
and 'int'
```

This might have lead some to invent unit vectors in the momentum update (e.g., `p = p + vector(Fmag,0,0)*dt`).

SL4 – Student created a new variable for \vec{p}_f . In computational modeling, the equal sign in a update line means “add and replace”. Some students used a new symbol for the final momentum (e.g. `pfinal`) and then replaced the momentum in the next step (e.g. `p = pfinal`). Others only did the former, that is, they did not replace the momentum with its updated value.

4. Other Codes

Two common errors were not included in the above codes because they do not reflect errors in the procedure of modeling the motion of the low-mass particle. We present two miscellaneous codes which were common enough to consider relevant.

O1 – Student attempted to update force, momentum or position for the massive particle. The massive particle was intended to remain in place.

O2 – Student did not attempt the problem. Some students uploaded plain text files to the receive bonus credit for uploading their code. We assumed they did not attempt solving the problem.

* Current Address: Department of Physics, University of Colorado at Boulder, Boulder, CO 80309

† Corresponding Author: michael.schatz@physics.gatech.edu

‡ Current Address: Advanced Instructional Systems, Inc., Raleigh, NC 27696

¹ K. Jordan and C. Woodward, *Siam statement on computational science and engineering*, <http://www.siam.org/students/resources/report.php>.

² F. Fenton, E. Cherry, A. Karma, and W. Rappel, *Modeling wave propagation in realistic heart geometries using the phase-field method*, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **15**, 013502 (2005).

³ J. González, U. Sperhake, and B. Brüggmann, *Black-hole binary simulations: The mass ratio 10 1*, *Physical Review D* **79**, 124006 (2009).

⁴ W. MacDonald, E. Redish, and J. Wilson, *The MUPPET manifesto*, *Computers in Physics* **2**, 23 (1988), ISSN 0894-1866.

⁵ E. Redish and J. Wilson, *Student programming in the introductory physics course: MUPPET*, *American Journal of Physics* **61**, 222 (1993).

⁶ H. Schecker, *Learning physics by making models*, *Physics Education* **28**, 102 (1993).

⁷ H. Schecker, in *Proceedings of the 1994 International System Dynamics Conference* (System Dynamics Society, 1994), pp. 74–84.

⁸ A. DiSessa and H. Abelson, *Boxer: a reconstructible computational medium*, *Communications of the ACM* **29**, 859 (1986), ISSN 0001-0782.

⁹ B. Sherin, A. diSessa, and D. Hammer, *Dynaturtle revisited: Learning physics through collaborative design of a*

- computer model, *Interactive Learning Environments* **3**, 91 (1993), ISSN 1049-4820.
- ¹⁰ N. Finkelstein, W. Adams, C. Keller, K. Perkins, and C. Wieman, *High-tech tools for teaching physics: The physics education technology project*, *Journal of Online Teaching and Learning* (2006).
 - ¹¹ R. Chabay and B. Sherwood, *Computational physics in the introductory calculus based course*, *American Journal of Physics* **76(4&5)**, 307 (2008).
 - ¹² R. Beichner, R. Chabay, and B. Sherwood, *Labs for the Matter & Interactions curriculum*, *American Journal of Physics* **78**, 456 (2010).
 - ¹³ B. A. Sherwood and R. Chabay, *Vpython – 3d programming for ordinary mortals*, <http://www.vpython.org/> (2011).
 - ¹⁴ W. Christian and A. Titus, *Developing web-based curricula using java physlets*, *Computers in Physics* **12**, 227 (1998).
 - ¹⁵ K. Perkins, W. Adams, M. Dubson, N. Finkelstein, S. Reid, C. Wieman, and R. LeMaster, *PhET: Interactive simulations for teaching and learning physics*, *The Physics Teacher* **44**, 18 (2006).
 - ¹⁶ C. Wieman and K. Perkins, *A powerful tool for teaching science*, *Nature physics* **2**, 290 (2006).
 - ¹⁷ F. Esquembre, *Easy java simulations: A software tool to create scientific simulations in java*, *Computer Physics Communications* **156**, 199 (2004).
 - ¹⁸ W. Christian and F. Esquembre, *Modeling Physics with Easy Java Simulations*, *The Physics Teacher* **45**, 475 (2007).
 - ¹⁹ M. Belloni and W. Christian (Presented at the 2007 AAPT Topical Conference: Computational Physics for Upper-Level Physics Programs, Davidson, NC, 2007).
 - ²⁰ F. Esquembre (Presented at the 2007 AAPT Topical Conference: Computational Physics for Upper-Level Physics Programs, Davidson, NC, 2007).
 - ²¹ D. Scherer, P. Dubois, and B. Sherwood, *Vpython: 3d interactive scientific graphics for students*, *Computing in Science & Engineering* **2**, 56 (2000).
 - ²² R. Chabay and B. Sherwood, *Matter and Interactions I: Modern Mechanics* (John Wiley and Sons, Inc., 2010), 3rd ed.
 - ²³ M. Kohlmyer, Ph.D. thesis, Carnegie Mellon University (2005).
 - ²⁴ We acknowledge that some courses might teach students to solve the problem of an falling object subjected to linear ($F \sim v$) air drag in one dimension analytically, but this hardly demonstrates the full predictive power of Newton's second law.
 - ²⁵ W. Conover, *Practical Nonparametric Statistics* (John Wiley and Sons, 1999).
 - ²⁶ The first administration of this assignment was during a regular hour exam. Roughly, 40% of the students modeled the motion correctly. However, students used their own laptop computers which created several logistical challenges.
 - ²⁷ P. Tan, M. Steinbach, and V. Kumar, *Introduction to data mining* (Pearson Addison Wesley Boston, 2006).
 - ²⁸ B. Everitt, S. Landau, and L. M., *Cluster Analysis* (Oxford University Press Inc., New York, NY, 2001).
 - ²⁹ R. Springuel, M. Wittmann, and J. Thompson, *Applying clustering to statistical analysis of student reasoning about two-dimensional kinematics*, *Physical Review Special Topics-Physics Education Research* **3**, 20107 (2007).
 - ³⁰ P. Jaccard, *Study of comparative distribution of flower in the porition of alpes and jura*, *Bulletin of Society of Natural Sciences Vaudoise* **37**, 241 (1901).
 - ³¹ R. Sokal and C. Michener, *A statistical method for evaluating systematic relationships*, *Multivariate statistical methods, among-groups covariation* p. 269 (1975).
 - ³² P. Sneath, *The application of computers to taxonomy*, *Microbiology* **17**, 201 (1957).
 - ³³ T. Sorenson, *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content*, *Kongelige Danske Videnskabernes Selskab* **5**, 1 (1948).
 - ³⁴ M. Schommer, *Effects of beliefs about the nature of knowledge on comprehension.*, *Journal of Educational Psychology* **82**, 498 (1990).
 - ³⁵ M. Schommer, *Epistemological development and academic performance among secondary students.*, *Journal of Educational psychology* **85**, 406 (1993).
 - ³⁶ A. Elby, *Helping physics students learn how to learn*, *American Journal of Physics* **69**, S54 (2001).
 - ³⁷ I. Halloun, in *International Conference on Undergraduate Physics Education. AIP Conference Proceedings* (American Institute of Physics, 1997), vol. 399, p. 605.
 - ³⁸ K. Perkins, W. Adams, S. Pollock, N. Finkelstein, and C. Wieman, in *2004 Physics Education Research Conference. AIP Conference Proceedings* (American Institute of Physics, 2005), vol. 790, p. 61.